图书情报工作
第 64 卷 第 5 期 2020 年 3 月

新型城镇化发展中农民工在线信息行为特征及演化研究*

■ 常颖[1,2] 王晰巍[1,3] 韦雅楠[1] 王铎[1]
[1]吉林大学管理学院 长春 130022 [2]长春财经学院 长春 130122
[3]吉林大学大数据管理研究中心 长春 130022

摘 要：[目的/意义]在我国加快推进新型城镇化战略的过程中，分析农民工群体的信息行为特征及演化，对实现精准信息扶贫和分析新型城镇化过程中农民工信息需求具有重要意义。[方法/过程]在对农民工进行界定的基础上，利用 Python 和 Xpath 路径语言，结合百度贴吧农民工数据，使用 Jieba 库实现分词操作并将数据按照时间和内容两个维度进行可视化研究，对农民工群体的在线信息行为特征及演化规律进行分析。[结果/结论]研究结论揭示在我国新型城镇化进程中，农民工用户在线信息获取行为中的获取内容异化程度不断加深，需求层次不断提高；在线信息发布行为时间特征具有政策导向性，政府出台的城镇化战略、支持农民工返乡、保障农民工权益等政策对农民的信息行为演进具有重要影响；在线信息共享行为主题特征呈现多元化。

关键词：新型城镇化 农民工 信息行为 特征 演化
分类号：G203
DOI：10.13266/j.issn.0252 – 3116.2020.05.004

1 引言

根据国家统计局 2019 年 2 月 28 日公布的最新统计数据，2018 年我国城镇人口占人口比重（城镇化率）为 59.58%，乡村常住人口较 2017 年末减少 1 260 万人；全国农民工总量 28 836 万人，比上年增加 184 万人，增长 0.6%。其中，本地农民工 11 570 万人，增长 0.9%；外出农民工 17 266 万人，增长 0.5%[1]。国家统计局 2019 年 4 月发布的《2018 年农民工监测调查报告》显示，农民工的主体为 1980 年及以后出生的新生代农民工，占全国农民工总量的 51.5%，大专及以上学历的农民工占 10.9%，进城农民工中可上网人数比例达到 92.1%[2]。在互联网快速发展及新型城镇化战略的推进过程中，农民工群体作为连接城市与农村二元结构的桥梁，因信息消费能力不断提升而成为我国信息经济发展中新的消费群体，其在线信息行为成为产业界和学术界关注的新问题。

国内外学者针对农民工这一特殊群体的信息需求特征展开了相关研究。国外学者 A. Sánchez-Soto 通过对墨西哥龙舌兰农民进行分类对比研究，发现农民信息的需求在于获得技能和知识，以改善他们的农场经营状况[3]；F. Aldosari 等采用随机抽样的方法选取了 183 名巴基斯坦农民作为调查对象，发现 78.7% 的受访者认为互联网是重要的农业信息来源，受访者的教育程度与接收的资料的应用情况有高度显著的关系；受访者的农业经验与接收到的资讯的应用并无显著关系[4]。国内学者刁松龄将农民工分为融入型、徘徊型和返乡型，其关注的信息分别是法律维权信息、文化娱乐信息和技能培训信息[5]；孙红蕾等认为农民工面临的住房、就业、医疗等问题离不开信息支持，这些信息需求的满足可以有效地加快城市融入的进程[6]；任婕通过研究发现，农民工群体在获取政府信息时，主要通过网上查询和询问亲朋好友或熟人，面临的主要困难是获取的信息无法解决实际问题[7]。

从现有研究成果来看，随着我国乡村振兴发展步伐的推进，农民工群体的信息需求特 征分析已经成为国内近几年关注的新问题。但是，现有的国内外研究成果多侧重于农民工群体的信息需求和信息获取渠道

* 本文系国家自然科学基金面上项目“信息生态视角下新媒体信息消费行为机理及服务模式创新研究”（项目编号：71673108）研究成果之一。
作者简介：常颖（ORCID：0000 – 0002 – 2994 – 6727），副教授，博士研究生，E-mail：84337244@qq.com；王晰巍（ORCID：0000 – 0002 – 5850 – 0126），大数据管理研究中心主任，教授，博士生导师；韦雅楠（ORCID：0000 – 0002 – 7416 – 2403），博士研究生；王铎（ORCID：0000 – 0002 – 5060 – 7893），博士研究生。
收稿日期：2019 – 10 – 30 修回日期：2020 – 01 – 18 本文起止页码：32 – 40 本文责任编辑：王传清

研究，基于我国发展的现实情境对新型城镇化发展过程中农民工在线信息行为特征及演化规律的研究还相对较少。本文试图解决以下问题：①通过公开数据挖掘，分析农民工在新型城镇化发展过程中在线信息获取行为的内容演化特征及信息需求演变特征是什么？②新型城镇化发展过程中农民工在线信息发布行为时间特征及随着中国新型城镇化发展呈现出的演进规律是什么？③能否从农民工历史性的面板数据分析农民工信息共享行为主题特征的演进规律，这些主题特征能否反映我国乡村振兴中农民工生活状态和思想的演变？本文通过百度贴吧中的“农民工吧”采集公开数据，并针对用户发帖的内容、发帖的时间及回帖的主题进行分析，进而剖析我国新型城镇化过程中农民工在线信息获取、信息发布和信息共享行为特征与规律，为提升面向农民工的信息服务水平、推动新型城镇化发展进程起到积极的作用。

2 相关研究

2.1 农民工的界定

1983 年首次有人提出“农民工”的概念，认为其是从事工人职业但具有农民身份的群体。1984 年 6 月，“农民工”这个词首次出现在中国社会科学院出版的《社会学通讯》中。2002 年 11 月，中共十六大报告提出“走中国特色城镇化道路”之后，农民工这一群体受到广泛关注。学者们从不同角度对农民工进行了界定。周大鸣认为界定农民工的关键为户口在农村，离开所在地兼有农民和打工者身份，从事非农业生产活动的群体[8]；卢海元将农民工界定为拥有农村土地的相关权利，以城市打工为生活来源，具有非城镇居民身份的农民群体[9]；邹晓美认为农民工是户籍保留在农村，但大部分时间不从事农业生产活动，以在城镇打工谋生的特定人群[10]。

在对相关研究进行梳理的基础上，结合我国在新型城镇化发展进程中农民工群体特征，本文将“农民工”的概念归纳为：在中国新型城镇化发展的进程中，由于区域经济发展失衡、求学、创业等原因形成的流动性人口，他们兼有农民与工人的两重身份，在农村具有土地承包经营权，以在城镇务工为主要生活来源，生活工作具有流动性，是主要从事第二、第三产业工作的群体。

2.2 农民工在线信息行为

信息行为指通过一系列文化层面和社会层面的手段，识别、搜寻、使用和分享来自报刊、电视、互联网等来源的信息的过程，一般由信息搜寻、信息使用和信息共享三部分组成[11]。国外学者 T. D. Wilson 认为信息行为是用户在某种信息需求的基础上进行的与信息渠道和信息资源有关的行为[12]；A. Spink 等建立的信息行为框架包括信息搜寻行为、信息觅食行为、信息共享行为和信息使用行为 4 个过程[13]；K. E. Pettigrew 等将信息行为界定为用户在不同的场景中需求、搜索、提供和利用信息的行为总和[14]。国内普遍认同胡昌平教授提出的观点，即信息行为是主体为了满足特定的信息需求，在外部条件的刺激下表现出的行为，包括信息获取、查找、交流、传播和利用等行为[15]。

本文在综合上述学者观点的基础上，结合本文的研究目的与研究群体特点，将农民工在线信息行为界定为：农民工在外部环境条件和自身发展的刺激下，通过互联网、移动互联网和移动终端等媒介，满足自身的信息需求，进行信息获取、信息发布和信息共享的行为总和。

2.3 新型城镇化发展中农民工在线信息行为特征问题的提出

2013 年 12 月，中央城镇化工作会议分析了我国城镇化发展的形势，讨论了《国家新型城镇化规划》，提出了推进新型城镇化的战略部署，明确形成基本公共服务和基础设施一体化、网络化发展的城镇化新格局。国务院 2017 年 8 月印发的《关于进一步扩大和升级信息消费持续释放内需潜力的指导意见》明确指出：“进一步扩大和升级信息消费，通过多种方式开展宣传引导活动，面向各类消费主体特别是信息知识相对薄弱的群体，普及信息应用、网络支付、风险甄别等相关知识；支持企业推广面向低收入人群的经济适用的智能手机、数字电视等信息终端设备，推介适合农村及偏远地区的移动应用软件和移动智能终端。”

农民工作为新型城镇化战略的主体，在互联网海量信息的环境下，农民工群体如何通过互联网、移动互联网、智能移动终端，进行信息获取、发布和共享，满足自身在求职、培训、权益保障、社交、返乡创业、城市融入等方面的成长和生活要求，通过自身信息素养的提升，最终完成新型城镇化发展过程中的“新农人”的社会转型，需要分析和了解农民工在线信息发布行为的内容特征、获取行为时间特征和共享行为的主题特征及演化规律，进而根据农民工在线信息行为的特点，提供更有针对性的农民工在线信息服务，解决国家在新型城镇化发展过程中“数字扶贫”问题。

3 研究设计

3.1 数据源选择

百度贴吧自 2003 年 11 月 26 日创建以来,基于搜索引擎建立的在线交流平台吸引了大量共同兴趣爱好的群体,是全球最大的中文交流社区。社会化媒体传播公司 We are social 发布的《2018 年全球数字报告》中指出,截至 2018 年 12 月,百度贴吧注册用户人数超过 10 亿,在线活跃用户高达 3 亿[16],InMobi 与 AdMaster 共同发布《2018 中国移动互联网用户行为洞察报告》中指出,百度贴吧在兴趣类 APP 的行业渗透率居于首位[17]。根据联通大数据沃指数移动应用 APP 排行榜显示,截至 2019 年 3 月,百度贴吧在社区论坛类排名第二位。百度贴吧所拥有的用户群体和文本信息蕴藏着庞大的挖掘价值,其信息环境具有开放性、传播方式具有多样性、贴吧功能具有创新性,为用户的信息获取提供了有利条件,为用户在线信息行为特征研究提供了大量可靠和客观的样本来源。

百度贴吧中"农民工吧"作为目前国内最活跃的农民工在线交流社区平台,自 2013 年 2 月创建至今,关注数为 17 745 人,贴吧主题数为 10 224 个,贴吧的封闭性保证了农民工群体交流的深入性。百度贴吧应用其在搜索引擎中的知名度,可以为农民工群体提供便捷的搜索同类型用户的途径。百度贴吧数据的时间序列特征,为农民工群体信息行为特征的演化分析提供了可靠的数据来源。因此本文将百度贴吧作为获取农民工在线信息行为的数据采集源,以阶段性面板数据来呈现和分析中国新型城镇化发展中农民工在线信息行为的特征及演化规律。

3.2 数据采集

本研究结合 Python 语言与 Xpath 即 XML 路径语言,通过 Google chrome 中开发者工具功能,查看元素 Element,在 URL 的页面选中某一元素 copy 其中的 Xpath,取到某元素的 Xpath,获得文本信息以及所在路径,利用 URL 中的 get 方法,根据不同的页码进行循环遍历,实现海量农民工数据的爬取,并将采集后的数据存入 PostgreSQL 数据库。采集数据内容包括"农民工吧"的贴吧信息和用户发帖信息。贴吧信息包括贴吧名称、贴吧主页 URL 地址、贴吧话题类型、关注者人数(会员数目)、贴吧帖子总数、贴吧标语、贴吧所在目录以及该贴吧的友情贴吧名称;用户发帖信息包括帖子标题、帖子与回帖内容、帖子所在主页的 URL 地址、帖子创建时间、帖子被回复数、帖子所属贴吧名、发帖人昵称、发帖人主页 URL 地址以及发帖人等级。

3.3 数据清洗

本文采集了 2013 年 2 月 1 日至 2019 年 2 月 1 日共 6 年的样本数据,共计 87 320 条。为了保证研究的客观性和代表性,对获得的数据进行清洗,包括检测与处理重复值、检测与处理数据的缺失值、检测与处理数据的异常值。在数据的清洗过程中,采用 Python 语言的 Pandas 数据分析包进行清洗。重复值包含记录重复与特征重复两种,记录重复采用列表(list)和集合(set)元素唯一性的特征去重;特征重复利用 Pandas 中相似度计算方法 corr,利用特征间相似度为 1 的特征去除重复数据。对于不完整数据(即值缺失),采用手工填入的方法。对于异常值,使用正则表达式剔除发帖内容中的空格、图片和表情;针对特殊信息比如广告、视频等,通过异常捕获的方式加以过滤。清洗后最终获得有效数据 55 450 条。

3.4 数据处理

本文数据处理的步骤如下:①对爬取的帖子和回帖内容信息进行归类分词。需要利用 Python 语言中的 Jieba 库实现分词操作,对于专有名词,手动更新词汇库,为每个关键字设置初始值,通过循环的方式,计算各个关键字出现的次数,并在控制台进行输出;②使用数据分析工具 IBM SPSS Modeler 中的 Text Analytics Module 模块将采集到的百度贴吧中"农民工吧"的内容进行分类,并从中提取关键词以及每类内容所占比例;然后使用 Category Web 功能对每个主题之间的相互关系进行可视化,从而更好地理解数据分类结果之间的相互关系;③基于日期、时间、词频统计农民工发帖信息,其中日期按照年份、月份、星期进行分类;时间以每两个小时为基准进行分类;词频分别按照年份以及互动信息进行分类,统计农民工在线信息行为的特征及演化规律。

4 数据结果分析

本文通过百度贴吧中"农民工吧"采集公开数据,通过对用户发帖的内容、发帖的时间及回帖的主题进行分析,进而剖析新型城镇化过程中农民工在线信息获取行为的内容特征、信息发布行为的时间特征和信息共享行为主题特征与演化规律。

4.1 农民工在线信息获取行为内容特征分析

通过对"农民工吧"中帖子内容信息进行提取,采用 TF-IDF 统计方法,逐个计算词项的 TF-IDF 值,选取前 200 个 TF-IDF 值最大的词代表该贴吧信息发布行

为内容特征的关键热词，分析结果见图1。数据分析结果表明，关键词热度前4位的分别是“农民工”“工作”“讨薪”和“交友”。农民工发帖过程中群体标签意识很强，“工作”和“讨薪”除“农民工”自身这个热词外出现的频率基本相同，说明近几年在新型城镇化发展过程中，农民工群体有很多人进城务工，但在求职和工作过程中面临很多的困境，同时还要承担薪资拖欠、合法权益得不到保障的生活和社会压力。同时，随着互联网的普及这一群体的社交需求也日益凸显，交友成为发帖的第四大主题。其他的热词频率差异不明显，表明农民工群体的在线信息内容相对还是较为分散。



图1　贴吧发帖热词分布

为分析农民工在不同阶段信息获取内容的演化规律，对上述热词按发帖时间进行梳理和分析，发现农民工群体发帖的热词内容在不同年度差异较大(见图2)。这些年度热词的演化体现出农民工在中国新型城镇化发展过程中信息获取需求的演变。2013年发帖的热词里出现了“子女”“希望”“微信”“培训”“招聘”“安全”“待遇”和“工伤”，这些热词体现了农民工对求职、家庭、工作和自身前途的担忧。2014年发帖的热词里出现了“高薪”“出国”“诚心诚意”“维权”“生死”“跪求”和“骗子”，这些热词体现出农民工进入城市工作后面临很多困境和无奈的心态，并开始有了维权的意识和对权益保障方面的信息获取需求，同时也开始有农民工为了寻求更高的收入出国，希望到其他国家工作来谋求更好的出路。2015年到2016年发帖的热词里出现“讨薪”“血汗钱”“黑心”“讨薪路”“拖欠”“政府、保障”“过年、返乡”等热词，可以看出农民工意识到政府对维护自身权益起到重要作用，并有进城务工后回家过年的思乡情节。2017年农民工关注的热点中开始出现“创业”这样的热词，看出国家出台的“大众创业”的政策在一定程度上也给农民返乡创业带来了新的热情和机遇。

4.2　农民工在线信息发布行为时间特征分析

通过“周”时间段和“小时”时间段对“农民工吧”的发帖量进行统计分析，数据统计结果见图3。图中不同颜色的面积代表了周一至周日贴吧的发帖量，从结果可以看出农民工日发帖量波动性较小，周六、周日的发帖量与其他日期相比无明显变化。这主要是由于农民工从事的多为第二、第三产业的工作，工作时间较长，与其他群体相比每个月休息的时间一般为2－3天，周末多为工作时间，这在一定程度上反映了农民工的工作生活特点及信息发布行为在“24小时”内的时间分布规律。统计贴吧里的发帖数量可以发现，发帖量的峰值主要集中在9:00－10:00、19:00－20:00和23:00－0:00这3个关键时间段。由于农民工群体所从事的行业特点，23:00－0:00为下班休息时间，上午9:00－10:00为上班前的休息时间，而19:00－20:00则是晚饭时间，这3个时间段农民工时间充裕、信息发布行为较为活跃，因此发帖量与其他时间段相比波动较大，其他时间段的发帖量相对平均，表明农民工在线信息发布行为时间相对较为集中。

按照月份和年份统计发帖量可以看出随着我国城镇化进程的推进农民工群体的发帖规律，数据统计结果见图4。随着互联网的普及，尤其是智能手机上网成本的不断下调，自2013年起“农民工吧”的发帖量呈现逐年增加的趋势。2016年和2017年发帖量出现高峰，从不同月份的统计情况来看，全年的发帖量通常在1－2月份春节前后和5－6月份两个时间段较为集中，会出现发帖量高峰。通过对这两个时间段发帖主题的统计，求职、招聘、薪资为主题的帖子数量占到发帖总数的67.4%，表明农民工群体在这一时间段内更换工作相对较为频繁，对于求职方面的信息需求相对较大。

通过年度农民工在线信息发布行为数据分析结果可以看出，随着互联网的普及和农民工信息素养的提升，农民工发帖的数量总体呈逐年上升的趋势。受国家政策引导和时代关注的变化，发帖主题在不同的时间段不断演进，其中，“拖欠工资”主题的发帖量2013年2月为486个，在2016年3－4月达到峰值，此后逐

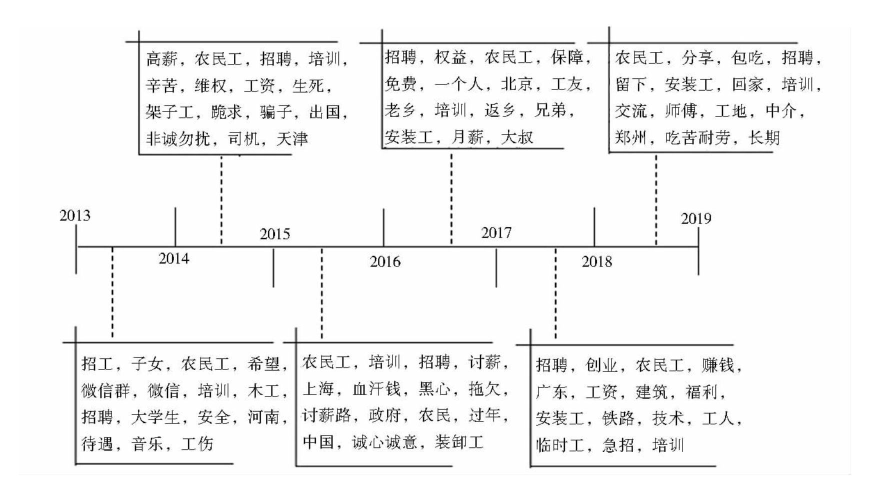


图 2　2013－2018 年农民工发帖热词演变

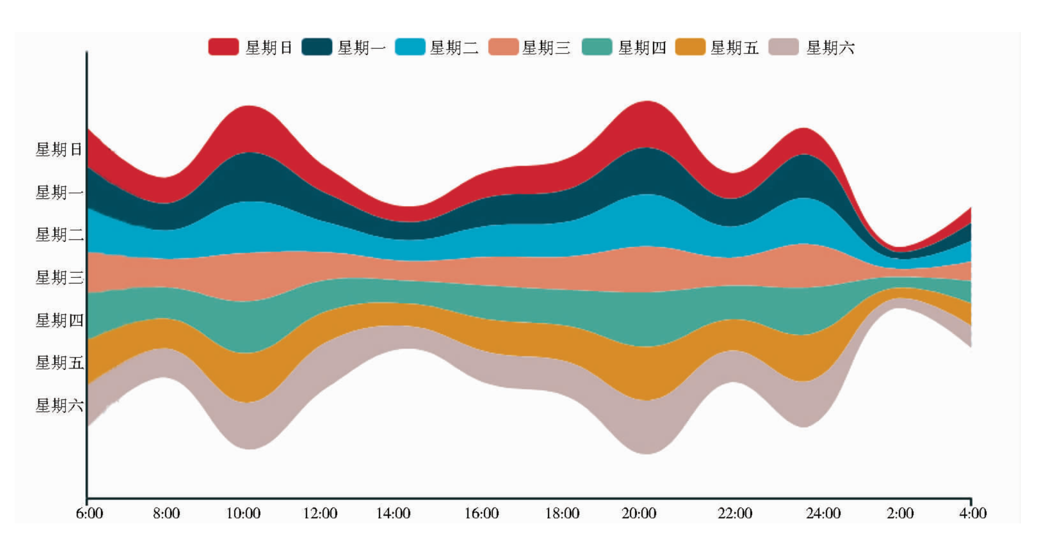


图 3　发帖量短期分布统计特征

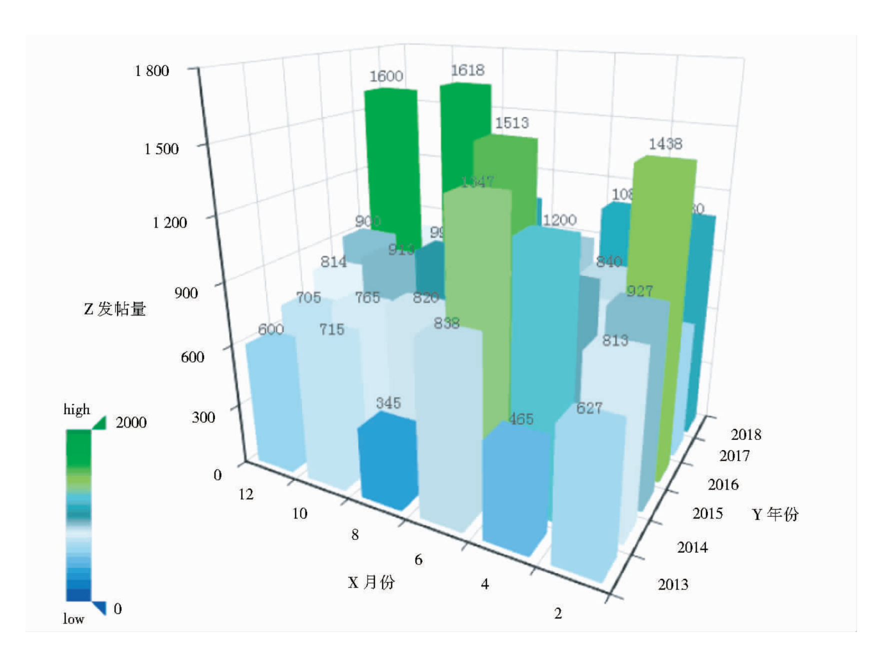


图 4　发帖量长期分布统计特征

年呈下降趋势，2018 年末回落到 302 个，表明“拖欠工资”问题一直是农民工群体密切关注的主题，但随着国家各类政策的不断出台，这一问题得到了切实的解决，其在农民工群体信息发布行为中所占的比重也逐年下降。此外，“权益维护”“社会保障”“返乡创业”“法律援助”主题的发帖量都呈现波动式上升趋势。2013 年 3－4 月这 5 类发帖量差异不明显，随着时间的推移，2018 年末发帖量中“返乡创业”达到 648 个、“法律援助”达到 613 个、“社会保障”达到 530 个、“权益维护”达到 433 个、“拖欠工资”达到 302 个，这在一定程度上也反映出农民工群体对国家政策的关注度在逐年提升，国家在新型城镇化发展过程中的很多政策对引导农民工群体的发展和推动城镇化发展起到了重要作用。

4.3 农民工在线信息共享行为主题特征分析

农民工群体在线信息共享行为的主题特征见图 5。根据之前分析的热词词频，以用户的层次关系为单向链，主题之间的关系相似度使用最短路径法，将所有的主题之间的关系表示为一个有向图。图 5 中每一个圆点代表一个主题的帖子数量，圆点的大小代表该主题帖子数量的多少。节点用于区分不同主题的帖子，父节点（发帖方）与子节点（回帖方）之间的连线表示针对同一个帖子的回复行为。连线的数量越多，表明用户之间对该主题的兴趣相似度越大。数据分析结果表明，信息共享较高的主题分别是“求职”“交友”“政府”“薪资”“讨薪”“培训”“子女教育”。其中“求职”“交友”“政府”“薪资”4 个主题之间的回帖率和共同关注度反应出农民工的信息共享需求整体相对较为集中，在线信息的共享主题特征反映出农民工群体解决日常工作和生活问题方式的重要转变。

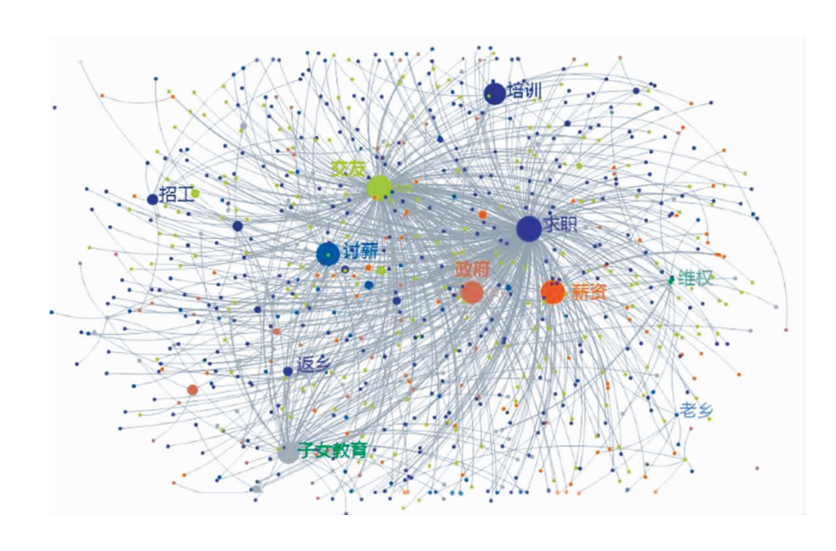


图 5　信息共享主题分布特征

5　讨论分析

5.1　农民工在线信息获取行为的内容特征规律

农民工在线信息获取行为的发帖热词年度演化数据结果表明，农民工群体在新型城镇化战略推进的过程中，在线信息获取行为中的获取内容异化程度不断加深，农民工信息获取需求层次不断提高。发帖的热词逐渐从满足生理需要的“招聘”“工资”演化为满足安全和情感需要的“权益”“保障”，再到满足尊重与自我实现需要的“微信群”“讨薪路”“创业”，这与魏巍等人的研究结果是吻合的[18]。农民工群体从农村来到城市，与原来的社会网络互动减少，需要时间适应并融入新的环境，在情感方面缺乏宣泄的机会和场所，因此“交友”成为农民工重要的在线信息获取需求，这与闫雪等人的研究结论是一致的[19]。

农民工在线信息获取行为的这一特征，也意味着针对这一群体提供信息服务时，各类信息服务机构应针对农民工群体建立信息服务平台，根据农民工不同的信息获取需求提供所需信息，帮助其解决工作和生活中遇到的实际问题。同时，应该成立专门的农民工服务平台，专门负责收集农民工的相关信息，对工作人员进行专业培训，面向各类型农民工提供个性化的信息服务，从而在更好地满足农民工信息获取需求的过程中，推动农民工整体素质提升，进而加快我国乡村振

兴的步伐。

5.2 农民工在线信息发布行为的时间特征规律

从农民工在线信息发布行为的演化规律来看，贴吧中相关信息发布行为特征在一定程度上可以反映出农民工在国家政策的导向下其心态和行为的演化轨迹。农民工群体由于城乡和区域经济发展的不均衡，在中国城镇化快速发展的过程中呈现了从农村迁徙到城市，然后又回乡创业的行为演化规律。因此，在发展中需要了解新的生存环境，解决居住、就业、培训、权益维护、市民化等系列问题。农民工在线信息行为的数据分析结果，根据"小时"时段和"周"时段的发帖量统计数据，农民工在线信息行为主要呈现 24 小时内的阶段性集中分布，即 9:00 – 10:00、19:00 – 20:00 和 23:00 – 0:00 这 3 个时间段，这与大学生群体在线信息活动时间 12:00 – 13:00、18:00 – 19:00 和 21:00 – 22:00存在较大差异[20]，与刘瑞芳开展的新生代农民工阅读推广服务研究结果相符合[21]。根据"月份"和"年份"的长期发帖规律统计分析可以看出，政府在出台城镇化发展战略、支持返乡创业、保障农民工合法权益等相关文件时，贴吧的发帖量变化较大，说明政府的政策导向对农民工的信息发布行为产生影响显著，这与汪汉清等人的研究结论也是一致的[22]。

针对农民工群体在线信息发布行为时间具有显著政策导向性这一特征，政府网站应加大对于农民工相关政策法规的宣传力度，以图片、短视频等农民工易于接受的方式进行宣传和解读，并根据农民工信息活动时间进行有针对性地推送。公共图书馆与社区平台应针对农民工居住较为集中的地域和群体提供在线政策法规咨询和解答服务，帮助农民工群体深入了解与自身密切相关的政策信息，切实解决中国城镇化发展过程中农民工的现实困难，进而提升国家乡村振兴、新型城镇化战略等政策的实施效果。

5.3 农民工在线信息共享行为的主题特征规律

在移动社交环境中，用户发布、评论、转发的内容多与自身感兴趣的主题密切相关，而由兴趣相似的群体共享的话题也具有同质性，这些用户在信息共享的过程中也会形成同一主题的虚拟社区[23]。百度贴吧中的"农民工吧"的数据分析结果表明，用户对于"求职""交友""政府""薪资"等与自身工作、生活密切相关的主题的关注度较高，对于共享信息的内容呈现多元化的需求，这与耿瑞利等人的研究结论是相符的[24]。用户对于共享信息需求层次具有差异，这一分析结果与杨绚等人通过对京籍农民工的田野调查研究结果相一致，即这一群体关注"必须知道的信息"，对"知道更好的信息"和"不需要被告知的信息"则关注度很低[25]。农民工在中国城镇化发展过程中，其就业和生活方面遇到的问题需要整合多源化的信息资源，由于农民工群体的信息意识、信息运用与分析能力相对低下[26]，农村相关信息服务配套设施不健全[27]，因此与其他群体比较而言农民工在线信息共享行为主题特征呈现多元化。这一结果与吴诗贤等人的研究一致[28]。

由于我国产业全球价值链提升，农民工非自愿性失业风险增加[29]；随着新型城镇化和精准扶贫工作的推进，农民群体的现代职业教育加快发展，国家也出台了相关政策鼓励应届农村高中生和农民工等报考[30]。由于 1980 年及以后出生的农民工数量已经占总人数的 51.5%，未来"职业教育""养老"和"医疗"等主题将成为我国新型城镇化发展中农民工群体关注的新焦点。因此，用工所在地政府部门可以构建由当地政府、用人企业和社区组成的联动信息综合服务平台，通过整合多元化信息渠道和信息资源，为辖区内的农民工提供招工求职、子女教育、权益维护、交友、养老和医疗等综合信息。同时，政府或行业机构等服务、中介机构应注意通过在线培训等方式提高农民工群体的信息整合能力，在政务信息、生活信息等推送的过程中重视信息所涵盖的真实性；并注意通过对社交网络中农民工群体中具有影响力的"大 V"农民工的识别与培训，做好农民工这一新兴网络舆情群体的监督与舆论引导，让农民工群体更好地融入城市生活，为新时代的乡村振兴和农民创业提供更好的服务和支持。

6 研究结论

本研究的理论价值在于通过对农民工群体在线信息行为的分析，通过百度贴吧中"农民工吧"的公开数据，从农民工在线信息获取行为的内容特征、发布行为的时间特征和共享行为的主题特征 3 个方面，对中国城镇化发展过程中农民工在线信息行为的特征和演化规律进行分析，在理论层面为我国农民工群体的信息行为分析提供理论框架。本研究的实践价值在于，通过分析近 6 年时间的面板数据，客观呈现中国城镇化发展过程农民工群体的在线信息行为的演化特征及演化规律。研究结果表明，在线信息获取行为中的获取

内容异化程度不断加深，农民工的信息获取需求层次不断提高；在线信息发布行为的时间特征随着我国城镇化不同发展阶段的政策导向而不断进行演化，政府的政策导向对农民工的信息发布行为产生较大影响；农民工在线信息共享行为主题特征呈现多元化，共享的主题多与自身工作、生活密切相关。

本研究也存在一定局限性。研究的数据来源仅为百度贴吧中的“农民工吧”，样本缺乏与其他群体贴吧的对比分析。在后续的研究中，将扩大数据来源的代表性，并增加与其他群体在线信息行为的比较研究。

参考文献：

[1] 国家统计局网站. 中国 2018 年国民经济和社会发展统计公报[EB/OL]. [2019－02－28]. http://www.gov.cn/shuju/2019-02/28/content_5369270.htm.

[2] 国家统计局网站. 2018 年农民工监测调查报告[EB/OL]. [2019－04－29]. http://www.stats.gov.cn/tjsj/zxfb/201904/t20190429_1662268.html.

[3] ARMANDO S S. Information needs and information behavior of blue agave farmers in Tequila, Jalisco: a case study [J]. Investigación Bibliotecológica: archivonomía, bibliotecología e información. 2016, 30(69): 137－169.

[4] FAHAD A, MOHAMED S. Farmers' perceptions regarding the use of information and communication technology (ICT) in Khyber Pakhtunkhwa, Northern Pakistan [J]. Journal of the Saudi Society of Agricultural Sciences. 2019, 18(2): 211－217.

[5] 刁松龄. 城市化进程中外来农民工信息服务研究——以珠三角为例[J]. 图书情报工作, 2009, 53(4): 136－139.

[6] 孙红蕾, 严昕, 郑建明. 基于环境扫描的城镇信息化发展战略转换研究[J]. 情报科学, 2017, 35(11): 23－27.

[7] 任婕. 农民工政府信息需求及获取渠道调查研究[J]. 图书馆研究与工作, 2018(9): 22－26.

[8] 周大鸣. 珠江三角洲人口移动探讨[J]. 社会科学战线, 1990(2): 148－152.

[9] 卢海元. 走进城市: 农民工的社会保障(前言)[M]. 北京: 经济管理出版社, 2004: 2.

[10] 邹晓美, 高泉. 农民工权利研究[M]. 北京: 中国经济出版社, 2010: 1.

[11] SAVOLAINEN R. Everyday information practices: a social phenomenological perspective[M]. Toronto: Scarecrow Press, 2008: 1－232.

[12] WILSONT D. Information behavior: and interdisciplinary perspective[J]. Information processing & management, 1997, 33(3): 551－572.

[13] AMANDA S, CHARLES C. New directions in human information behavior [M]. Berlin: Springer, 2006: 332.

[14] PETTIGREW K E, FIDEL R, BRUCE H. Conceptual frameworks in information behavior [J]. Annual review of information science and technology, 2001, 35: 43－78.

[15] 胡昌平. 现代信息管理机制研究[M]. 武汉: 武汉大学出版社, 2004: 124－132.

[16] 搜狐网. 2018 全球数字报告[EB/OL]. [2019－01－30]. http://www.100ec.cn/detail--6435296.html.

[17] 中国报告网. 2018 年移动互联网行业数据研究报告[EB/OL]. [2019－01－26]. http://news.chinabaogao.com/it/201901/0126396102019.html.

[18] 魏巍, 黄丽霞. 基于马斯洛需求层次理论的农民工信息需求分析[J]. 图书馆学研究, 2016(5): 58－62.

[19] 闫雪, 黄水清, 茆意宏. 发达地区农村外来劳动力的信息需求研究[J]. 图书情报工作, 2012, 56(12): 54－59.

[20] 严承希, 王军. 高校学生网络行为时序特征的可视化分析[J]. 情报学报, 2018, 37(9): 890－904.

[21] 刘瑞芳, 武晓丽. 公共图书馆开展新生代农民工阅读推广服务探析[J]. 图书馆工作与研究, 2013(11): 93－95.

[22] 汪汉清, 樊振佳. 我国涉农信息服务政策核心话语演变分析: 以 1980 年以来中央“一号文件”为例[J]. 图书情报工作, 2019, 63(4): 96－106.

[23] 张宁, 袁勤俭, 朱庆华. 个体认知专注与虚拟社区参与关系的元分析[J]. 情报学报, 2018, 37(2): 161－171.

[24] 耿瑞利, 申静. 不同文化视域下社交网络用户知识共享行为动机研究[J]. 中国图书馆学报, 2019(1): 60－81.

[25] 杨绚. 针对农民工的政府信息的传播渠道研究[D]. 北京: 中国社会科学院研究生院, 2014.

[26] 区晶莹, 高雅婧, 黄丽. 新市民信息能力对其城市融入的影响研究——以珠三角为例[J]. 图书情报工作, 2017, 61(8): 25－32.

[27] 李国新. 关于加强农村公共文化服务建设的思考[J]. 中国图书馆学报, 2019(7): 4－11.

[28] 吴诗贤, 张必兰. 农转城新市民信息素养与城市社会融合度的神经网络映射模型[J]. 图书情报工作, 2013, 57(23): 48－52.

[29] 北京师范大学劳动力市场研究中心. 2019 中国劳动力市场发展报告[EB/OL]. [2019－12－15]. http://labor.bnu.edu.cn/news/2019/1226/443.html.

[30] 中国政府网. 2019 年国务院政府工作报告[EB/OL]. [2019－03－05]. http://www.gov.cn/premier/2019-03/16/content_5374314.html.

作者贡献说明：

常颖：负责论文主体内容的撰写；
王晰巍：负责论文框架和论文的修改；
韦雅楠：论文英文内容的翻译及处理；
王铎：协助进行论文数据的收集和整理。

Research on the Characteristics and Evolution of Online Information Behavior of Migrant Workers in the Development of New-type Urbanization

Chang Ying[1,2]　Wang Xiwei[1,3]　Wei Yanan[1]　Wang Duo[1]
[1] School of Management, Jilin University, Changchun 130022
[2] Changchun University of Finance And Economics, Changchun 130122
[3] Big Data Management Research Center, Jilin University, Changchun 130022

Abstract: [Purpose/significance] In the process of accelerating the new urbanization strategy in China, it is of great significance to analyze the information behavior characteristics and evolution of migrant workers, so as to realize accurate information poverty alleviation and analyze the information needs of migrant workers in the new urbanization process. [Method/process] Based on defining the migrant workers, on the basis of using the Python language and Xpath with the baidu post bar workers data, using Jieba library to realize participle operation according to the two dimensions of time and content to the data visualization research, online information sharing for the migrant workers group behavior characteristics and evolution were analyzed. [Result/conclusion] The research conclusion reveals that in the process of new urbanization in China, the alienation degree of the acquired content in the online information acquisition behavior of migrant workers continues to deepen, and the level of demand continues to improve. The time feature of online information release behavior is policy-oriented. The urbanization strategy issued by the government, supporting the return of migrant workers to their hometowns, protecting the rights and interests of migrant workers and other policies have an important impact on the evolution of farmers' information behavior. The theme features of online information sharing are diversified.

Keywords: new-type urbanization　migrant worker　information behavior　characteristics　evolution

《图书情报工作》2020 年选题指南

[编者按]本选题指南是根据本刊的定位、性质与发展需要,结合图情档学科前沿热点及当前与未来需要解决的重要问题,邀请本刊编委和青年编委为本刊策划定制,再经编辑部整理、修改和补充而形成的。这是本刊 2020 年度关注、报道的重点领域(包括但不限于这些选题),供作者选题和研究以及向本刊投稿时的参考和借鉴。

1. 中国特色图情档学科体系、学术体系、话语体系建设
2. 图情档一级学科建设与融合发展战略
3. 图书馆"十四五"规划编制的关键问题
4. 国家文献信息资源保障能力及其建设
5. 开放科学背景下信息资源建设问题
6. 全民阅读中图书馆的定位与担当
7. 图书馆空间服务的理论与实践
8. 嵌入式学科服务的绩效评价与管理
9. 公众科学、科学素养与泛信息素养
10. 图书馆服务本科教育的模式与能力
11. 图书馆文化传承与文化育人的理论与实践
12. 图书馆出版与出版服务
13. 新媒体时代图书馆科学传播的功能与实践
14. 图书馆营销推广的战略与策略研究
15. 图书馆泛合作研究的实践与理论
16. 国家区域发展战略下图书馆联盟建设与创新服务
17. 网络空间治理的情报学问题
18. 知识产权信息服务能力与效果评估
19. 信息分析中的新技术与新方法
20. 情报服务标准化与评价
21. 数字人文与数字学术的研究与实践
22. 人工智能在图情档中的应用
23. 图书馆智能服务与智慧服务
24. 开放数据生态中的元数据发展模式研究
25. 开放科学数据行为及其模型构建
26. 数据资源建设与数据馆员能力建设
27. 大数据时代信息组织与知识组织
28. 科学数据管理与服务
29. 学术成果监测与学科竞争力分析
30. 情报计算(计算情报)的理论与方法
31. 情报分析服务质量与效能评价
32. 情报研究与智库研究的关系
33. 科学与技术前沿分析理论与方法
34. 健康中国 2030 战略下的健康信息学
35. 人机交互行为及服务模式创新
36. 图情档在新型智库建设中的作用机制
37. 智能信息服务的理论和方法
38. 数字公共文化资源、服务与体系建设
39. 数据时代政务信息资源管理和开发利用
40. 数字档案馆生态系统治理策略
41. 档案数据治理理论与治理体系
42. 政府数据开放平台应用与评价
43. 社会记忆视角下档案信息资源整理、保护与开发
44. 民族文献遗产产业化开发与利用
45. 图情档学科教育模式与人才培养能力